

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: SIMON PAUL DAVIS

Serial No.: 09/915,553 Group Art Unit:

Filed: JULY 27, 2001 Examiner:

Title: IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

CLAIM FOR PRIORITY UNDER 35 U.S.C. §119

Commissioner for Patents
Washington, D.C. 20231


Sir:

The benefit of the filing date of prior foreign applications numbers 0018328.5 and 0024468.1, filed in Germany on July 17, 2000 and October 6, 2000 respectively, is hereby requested and the right of priority under 35 U.S.C. §119 is hereby claimed.

In support of this claim, filed herewith is a certified copy of the original foreign application.

Respectfully submitted,

September 28, 2001



Gary R. Edwards
Registration No. 31,824

CROWELL & MORING, LLP
P.O. Box 14300
Washington, DC 20044-4300
Telephone No.: (202) 624-2500
Facsimile No.: (202) 628-5116

THIS PAGE BLANK (USPTO)



CERTIFIED COPY OF PRIORITY DOCUMENT

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

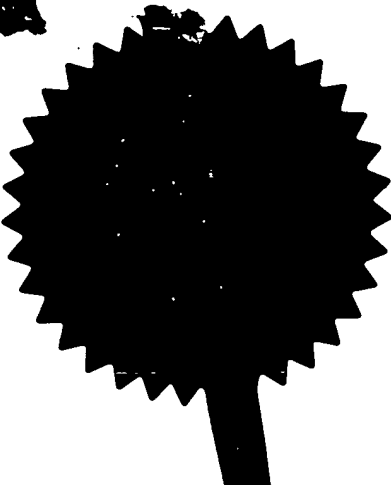
Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed



Dated

27 July 2001

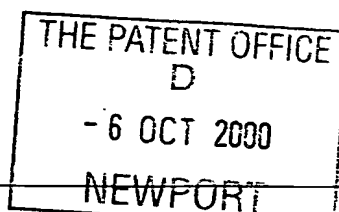


THIS PAGE BLANK (USPTO)

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

06 OCT 2000



060CT00 E573920-1 D00393
P01/7700 0.00-0024468.1

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

2000P04909/GB/R76/MM/rr

2. Patent application number

(The Patent Office will fill in this part)

0024468.1

3. Full name, address and postcode of the or of each applicant (underline all surnames)

ROKE MANOR RESEARCH LIMITED

OLD SALISBURY LANE
ROMSEY, HAMPSHIRE, SO51 0ZN
UNITED KINGDOM

Patents ADP number (if you know it) 56157455006

If the applicant is a corporate body, give the country/state of its incorporation

UNITED KINGDOM

4. Title of the invention

IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

5. Name of your agent (if you have one)

DEREK ALLEN

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Siemens Shared Services Limited
Intellectual Property Department
Siemens House, Oldbury
Bracknell, Berkshire RG12 8FZ
United Kingdom

Patents ADP number (if you know it) 7761000002

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number
(if you know it)

Date of filing
(day / month / year)

GB

0018328.5

27/07/2000

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

YES

- a) any applicant named in part 3 is not an inventor, or
 - b) there is an inventor who is not named as an applicant, or
 - c) any named applicant is a corporate body.
- See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	0
Description	9
Claim(s)	1
Abstract	1
Drawing(s)	3 57

57

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (Patents Form 7/77) 1 ✓

Request for preliminary examination and search (Patents Form 9/77) 1 ✓

Request for substantive examination (Patents Form 10/77)

Any other documents (please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Margaret Mackett

Date 4.10.2000

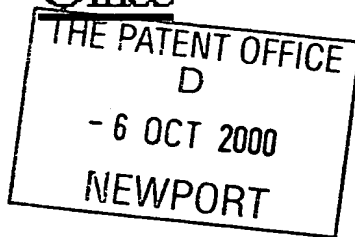
12. Name and daytime telephone number of person to contact in the United Kingdom Margaret Mackett 01344 396808

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.



**Statement of inventorship and of
right to grant of a patent**

06 OCT 2000

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

2000P04909/GB/R76/MM/rr

2. Patent application number
(If you, know it)

0024468.1

3. Full name of the or of each applicant

ROKE MANOR RESEARCH LIMITED

4. Title of the invention

IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

5. State how the applicant(s) derived the right
from the inventor(s) to be granted a patent

By provisions of contract of service and Section 39(1) (a) and (b) of
the Patent Act 1977.

6. How many, if any, additional Patents Forms
7/77 are attached to this form?
(see note (c))

ONE

I/We believe that the person(s) named over the page (*and on any
extra copies of this form*) is/are the inventor(s) of the invention
which the above patent application relates to.

Signature

Date 4.10.2000

Margaret Mackett

8. Name and daytime telephone number of
person to contact in the United Kingdom

Margaret Mackett

01344 396808

Notes

- a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- b) Write your answers in capital letters using black ink or you may type them.
- c) If there are more than three inventors, please give the names and addresses of the other inventors on the back of another Patents Form 7/77 and attach it to this form.
- d) When an application does not declare any priority, or declares priority from an earlier UK application, you must provide enough copies of this form so that the Patent Office can send one to each inventor who is not an applicant.
- e) Once you have filled in the form you must remember to sign and date it.

Patents Form 7/77

Enter the full names, addresses and postcodes of the inventors in the boxes and underline the surnames

DAVIS, Simon Paul
17 Westering
Romsey
Hampshire
SO51 7LX

Patents ADP number (if you know it): 587377 3002

Patents ADP number (if you know it):

Patents ADP number (if you know it):

Reminder

Have you signed the form?

Patents Form 7/77

IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

The present invention relates to improvements in or relating to switching devices, and is more particularly concerned with a method of adjusting bandwidth in such devices.

5 Traffic volume in the Internet is growing exponentially, doubling every three to six months. The current capacity of Internet Protocol (IP) routers is insufficient to meet this demand and hence products that can route IP traffic at extremely large aggregate bandwidths of the order of 10Gbit/s to several Tbit/s. Such routers are termed "Super Routers".

10 Additionally, there is a growing need to support multicast (one to many or many to many) communications within the internet or any other IP based network. To support such a service, an IP router must be able to replicate packets and send them to multiple outputs on a per packet basis. In a router where bandwidth allocations are strictly controlled (in order to
15 support Quality of Service criteria), it is necessary to determine how much bandwidth to allocate to multicast traffic across the core switching fabric.

It is therefore an object of the present invention to provide a method which overcomes the problems mentioned above.

In accordance with one aspect of the present invention, there is
20 provided a method of allocating bandwidth for multicast traffic in a switching device connected between a plurality of ingress means and a plurality of egress means, the method comprising the steps of:-

- a) determining ingress multicast queue occupancy for each ingress means;
- 25 b) determining the number of multicast cells received by the egress means from the ingress means in the last bandwidth allocation period;

c) determining the bandwidth at each ingress means and egress means after real time bandwidth allocations have been taken into account; and

5 d) calculating the bandwidth allocation for the next bandwidth allocation period in accordance with the values determined in steps a), b) and c).

For a better understanding of the present invention, reference will now be made, by way of example only, to the accompanying drawings in which:-

10 Figure 1 illustrates an ingress forwarder scheduling function;

Figure 2 illustrates centralised bandwidth allocation in accordance with the present invention; and

Figure 3 illustrates effective queue lengths for non-real time multicast bandwidth allocation.

15 The present invention relates to a method of dynamically adjusting the bandwidth, allocated to multicast traffic, across an asynchronous transfer mode (ATM) switch or crossbar like switching fabric that joins several IP packet forwarder functions to form a "Super Router" node.

20 In order to prevent head of line blocking, unicast traffic is queued in separate logical scheduling entities (called scheduler blocks) according to which egress forwarder it is destined. The scheduler block serves a set of queues (per class or per connection) via any mechanism desired (e.g. strict priority or Weighted Fair Queuing) provided that the real time IP traffic class is guaranteed a minimum bandwidth.

25 However, for multicast traffic, it is not practical to queue traffic on the basis of a unique combination of egress destinations. This is because the number of queues required becomes unmanageable even for a relatively

small number of egress ports. Hence, a separate multicast scheduler block is used in each ingress forwarder containing one real time multicast queue and one or more non-real time multicast queues as shown in Figure 1.

Figure 1 shows an ingress forwarder 100 which includes a unicast scheduler block 110 for scheduling unicast traffic and a multicast scheduler block 130 for scheduling multicast traffic. Although only one unicast scheduler block 110 and one multicast scheduler block 130 are shown, it will be appreciated that any number of such scheduler blocks may be provided in any combination according to a particular application.

The unicast scheduler block 110 comprises a plurality of queues 112, 114, 116 connected to a scheduler 118 which has an output 120 connected to a particular egress forwarder (not shown), for example, egress forwarder 1 as indicated. Although only three queues 112, 114, 116 are shown, it will readily be understood that any number of queues may be provided in accordance with a particular application.

The scheduler 118 has a scheduling rate which is determined by unicast bandwidth allocation and operates to transmit cells 122, 124, 126 at the head of respective queues 112, 114, 116 according to their priority, as indicated by arrow 128, to the output 120. Unicast bandwidth allocation is described in co-pending British patent application no. 9907313.2 (docket number F21558/98P4863), incorporated herein by reference.

The multicast scheduler block 130 comprises queues 132, 134 – a real time queue 132 and a non-real time queue 134. Both queues 132, 134 are connected to a scheduler 136 through which all multicast traffic passes. The scheduler 136 has an output 138.

It will readily be appreciated that although only one real time queue and one non-real time queue are shown, there may any number of such queues depending on the particular application.

Cells 142, 144 at the head of respective ones of the queues 132, 134
5 are selected for passage through the scheduler 136 to output 138 in accordance with a priority basis as indicated by arrow 146.

Incoming IP traffic from the line is queued in the relevant queues associated with a scheduler block. The choice of scheduler block is governed by the destination egress forwarder and whether it is multicast or unicast
10 traffic. The class of service determines the specific queue to be utilised.

The overall view of a centralised bandwidth allocation arrangement 200 is shown in Figure 2. The arrangement 200 comprises a plurality of ingress forwarders 210, a plurality of egress forwarders 220, a switching network 230 and a bandwidth allocation controller 240. Each ingress
15 forwarder 212, 214, 216, 218 can be connected to one or more egress forwarders 222, 224, 226, 228 as required via the switching network 230 under the control of the bandwidth allocation controller 240.

Although only four ingress forwarders 212, 214, 216, 218 and four egress forwarders 222, 224, 226, 228 are shown, it will be appreciated that
20 any number of ingress and egress forwarders can be provided in accordance with a particular application.

As shown in Figure 2, each ingress forwarder 212, 214, 216, 218 interfaces with the bandwidth allocation controller 240 via links 242 and 244 – only the links 242, 244 to ingress forwarder 212 being shown for clarity.
25 Link 242 provides the bandwidth allocation controller 240 with information relating to buffer occupancy, arrival rate of packets of cells etc for each ingress forwarder 212, 214, 216, 218. Link 244 provides each ingress

forwarder 212, 214, 216, 218 with scheduling decisions made by the bandwidth allocation controller 240.

Similarly, each egress forwarder 222, 224, 226, 228 interfaces with the bandwidth allocation controller 240 via link 246 which provides the
5 bandwidth allocation controller 240 with information relating to the multicast cells sent. Again, only the link 246 with egress forwarder 222 is shown for clarity.

For every fixed period, that is, the Bandwidth Allocation Period (BAP), each ingress forwarder 212, 214, 216, 218 sends buffer occupancy
10 (and possibly other information) via link 242 to the central bandwidth allocation controller 240. In addition, each egress forwarder 222, 224, 226, 228 sends information via link 246 on how many multicast cells were received in the last BAP from each ingress forwarder 212, 214, 216, 218. The bandwidth allocation controller 240 works out the allocation of
15 bandwidth between all ingress/egress forwarder pairs for the next BAP and uses this to provide scheduling information to the ingress forwarders 212, 214, 216, 218 via link 244 telling them which cells/packets to transmit in the next cell period.

However, in order to include multicast functionality into the
20 bandwidth allocation process some additions are required to the unicast algorithm defined in British patent application no. 9907313.2 mentioned above. The unicast bandwidth allocation algorithm essentially divides the available bandwidth at ingress and egress amongst competing forwarders using the ingress queue length as a weighting factor. The queue length of the
25 unicast scheduler block for traffic on ingress forwarder i destined for egress forwarder j is denoted by q_{ij} . Thus, for example, the amount of bandwidth

forwarder 212, 214, 216, 218 with scheduling decisions made by the bandwidth allocation controller 240.

Similarly, each egress forwarder 222, 224, 226, 228 interfaces with the bandwidth allocation controller 240 via link 246 which provides the
5 bandwidth allocation controller 240 with information relating to the multicast cells sent. Again, only the link 246 with egress forwarder 222 is shown for clarity.

For every fixed period, that is, the Bandwidth Allocation Period (BAP), each ingress forwarder 212, 214, 216, 218 sends buffer occupancy
10 (and possibly other information) via link 242 to the central bandwidth allocation controller 240. In addition, each egress forwarder 222, 224, 226, 228 sends information via link 246 on how many multicast cells were received in the last BAP from each ingress forwarder 212, 214, 216, 218. The bandwidth allocation controller 240 works out the allocation of
15 bandwidth between all ingress/egress forwarder pairs for the next BAP and uses this to provide scheduling information to the ingress forwarders 212, 214, 216, 218 via link 244 telling them which cells/packets to transmit in the next cell period.

However, in order to include multicast functionality into the
20 bandwidth allocation process some additions are required to the unicast algorithm defined in British patent application no. 9907313.2 mentioned above. The unicast bandwidth allocation algorithm essentially divides the available bandwidth at ingress and egress amongst competing forwarders using the ingress queue length as a weighting factor. The queue length of the
25 unicast scheduler block for traffic on ingress forwarder i destined for egress forwarder j is denoted by q_{ij} . Thus, for example, the amount of bandwidth

allocated to unicast traffic from ingress forwarder *i* to egress forwarder *j* at the egress, be_{ij} , is given by the following equation:-

$$be_{ij} = ABWE_j * \frac{q_{ij}}{\sum_k q_{ik}} \quad (1)$$

Here, $ABWE_j$ is the available bandwidth after real time reservations have been accounted for at the egress forwarder *j*, and $\sum q_{ik}$ is the sum of the buffer occupancies for data destined for egress forwarder *j* in every ingress forwarder. The term q_{ik} represents the buffer occupancy of the scheduler block in ingress forwarder *i* destined for egress forwarder *j*.

For real time multicast flows, the fan-out and bandwidth guarantees are known in advance and the sum of all ingress and egress requirements can be subtracted from the available bandwidth in the same way as for real time unicast traffic flows.

As the amount of egress bandwidth required for non-real time multicast flows is not known (compared with the case for real time multicast), it must be determined by the system. One way of determining the amount of egress bandwidth required is to collect statistics at the egress forwarders on the number of multicast cells received from each ingress forwarder in the last BAP. These statistics can then be included in the queue statistics message sent from the ingress forwarder to the central scheduler every BAP.

Although Figures 1 and 2 have been described with reference to ingress and egress forwarders, it will be appreciated that the ingress and egress devices are not limited to such devices and may comprise any suitable device which enables packets of data to be transferred from one side of a switching network to another.

Figure 3 illustrates a system 300 for calculating non-real time multicast bandwidth allocation. The system 300 comprises a plurality of ingress forwarders or line interface cards (LICs) 310, a plurality of egress forwarders or line interface cards (LICs) 320, and a switching network 330.

5 Each ingress LIC 312, 314, 316 has at least one queue 342, 344, 346 as shown. Each egress LIC 322, 324, 326 receives data from one or more ingress LICs 312, 314, 316 across the switching network 330 as shown. Only one queue 342, 344, 346 is shown in each ingress LIC 312, 314, 316 for clarity.

10 The ingress forwarder multicast queue occupancy is denoted as mcq_i for ingress forwarder i . The number of multicast cells received by egress forwarder j from ingress forwarder i in the last BAP is denoted by mcq_{ij} . The bandwidth allocated to non-real time multicast flows from ingress forwarder i to egress forwarder j is denoted by mcb_{ij} .

15 The value of mcq_i is used in the ingress bandwidth fair share in the same manner as q_{ij} does in the unicast centralised bandwidth allocation algorithm.

The values mcq_{ij} take part in the egress fair share allocation by providing a proportion with which to scale the ingress multicast queue occupancies. This means that the effective weight that the occupancy of the ingress non-real time (nrt) multicast queue (mcq_i) has on an egress forwarder j (called $emcq_{ij}$) is determined by the proportion of nrt cells received by egress forwarder j compared to all those received at egress forwarder j in the last BAP period. It is therefore governed by the following equation:-

25
$$emcq_{ij} = mcq_i * \frac{mcq_{ij}}{\sum_k mcq_{ki}} \quad (2)$$

The value of $emcq_{ij}$ will be used in egress bandwidth allocation functions alongside the unicast equivalents q_{ij} .

Thus, the equivalent of equation (1) when including the multicast traffic is given in equation (3):-

$$5 \quad bme_{ij} = ABWE_j * \frac{emcq_{ij}}{\sum_k q_{kj} + \sum_k emcq_{kj}} \quad (3)$$

where bme_{ij} is the nrt multicast bandwidth allocated between ingress forwarder i and egress forwarder j , $ABWE_j$, $emcq_{ij}$ and q_{ik} are the same as before.

Similar principles can be applied at the ingress for bandwidth
10 allocation and any left over bandwidth can be distributed between unicast and multicast allocations by any fair mechanism required.

The ingress equation for nrt multicast bandwidth becomes:-

$$bmi_{ij} = ABWI_i * \frac{imcq_{ij}}{\sum_k q_{jk} + \sum_k imcq_{jk}} \quad (4)$$

$ABWI_i$ is the available bandwidth at the ingress forwarder i after real
15 time traffic reservations have been taken into account. The term $imcq_{ij}$ is the ingress equivalent of $emcq_{ij}$ and is the effective weight with which to scale the ingress multicast queue occupancy. It is calculated from equation (5).

$$imcq_{ij} = mcq_i * \frac{mcq_{ij}}{\sum_k mcq_{ik}} \quad (5)$$

The actual allocated multicast bandwidth between ingress forwarder i
20 and egress forwarder j is the minimum of the ingress and egress bandwidth allocations as defined in equation (6).

$$bm_{ij} = \min(bmi_{ij} \quad bme_{ij}) \quad (6)$$

Any remaining bandwidth not allocated after this process has been carried out can be distributed between unicast and multicast allocations in any fair manner desired.

The advantages of the scheme described above include:

- 5 a) A 100% efficient distribution of bandwidth
- b) A fair distribution of nrt bandwidth according to ingress queue occupancies
- c) Prevention of overload of the cross-bar by restricting the total amount of bandwidth allocated at ingress and egress ports to the
- 10 cross-bar.

CLAIMS:

1. A method of allocating bandwidth for multicast traffic in a switching device connected between a plurality of ingress means and a plurality of egress means, the method comprising the steps of:-

a) determining ingress multicast queue occupancy for each ingress means;

b) determining the number of multicast cells received by the egress means from the ingress means in the last bandwidth allocation period;

c) determining the bandwidth at each ingress means and egress means; and

d) calculating the bandwidth allocation for the next bandwidth allocation period in accordance with the values determined in steps a), b) and c).

2. A method of allocating bandwidth for multicast traffic substantially as hereinbefore described with reference to the accompanying drawings.

ABSTRACT
IMPROVEMENTS IN OR RELATING TO SWITCHING DEVICES

Described herein is a method for making bandwidth allocation for data to be sent from a plurality of ingress forwarders or LICs (310, 312, 314, 316) to a plurality of egress forwarders or LICs (320, 322, 324, 326) across a routing device (330). The data may be unicast and/or multicast. The bandwidth allocation is calculated in accordance with ingress forwarder multicast queue occupancy for each ingress forwarder (312, 314, 316), the number of multicast cells received by egress forwarders (322, 324, 326) from the ingress forwarders in the last bandwidth allocation period, and the bandwidth allocated to non-real time multicast flows from ingress forwarders (310, 312, 314, 316) to egress forwarders (320, 322, 324, 326).

(Fig. 3)

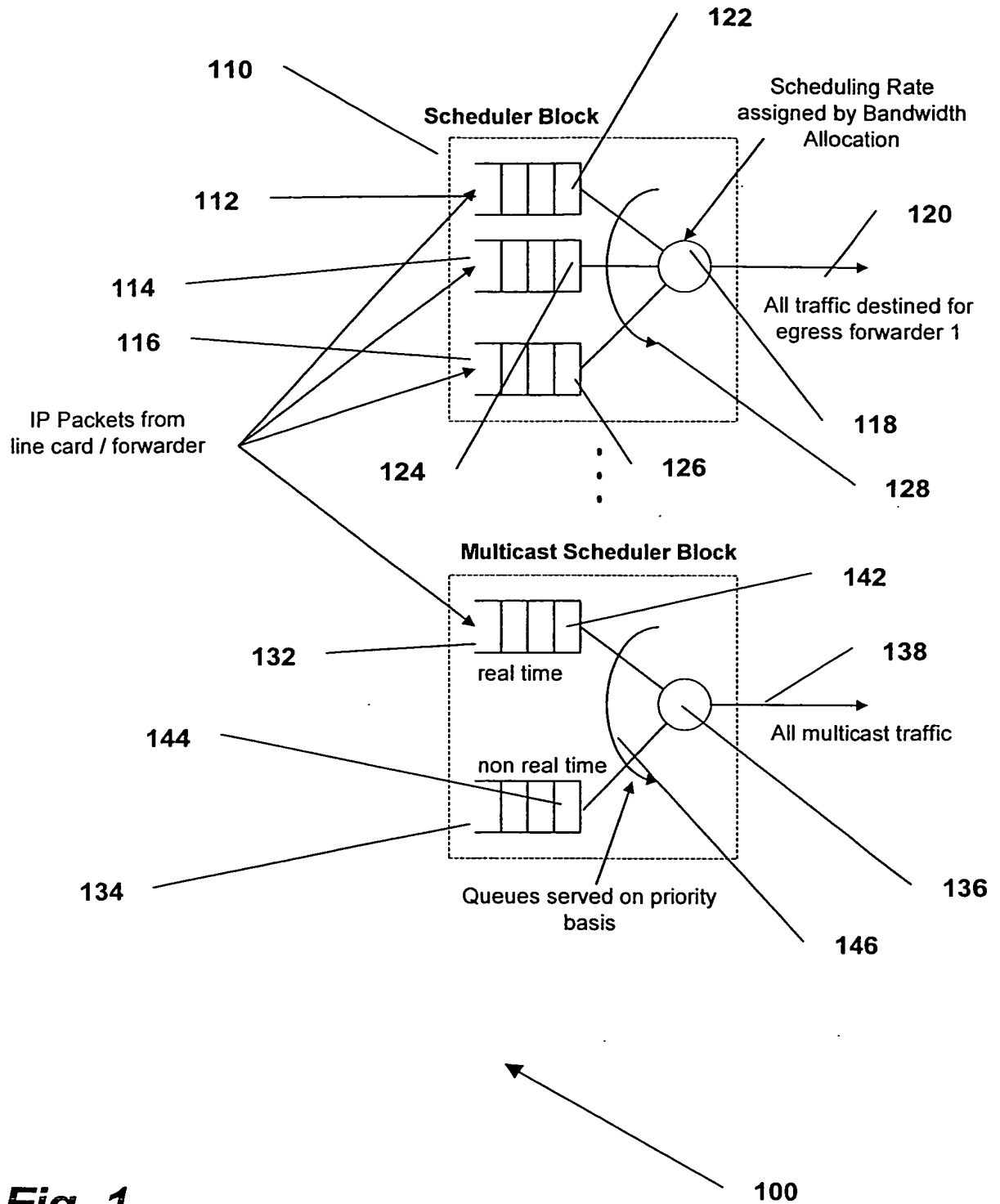
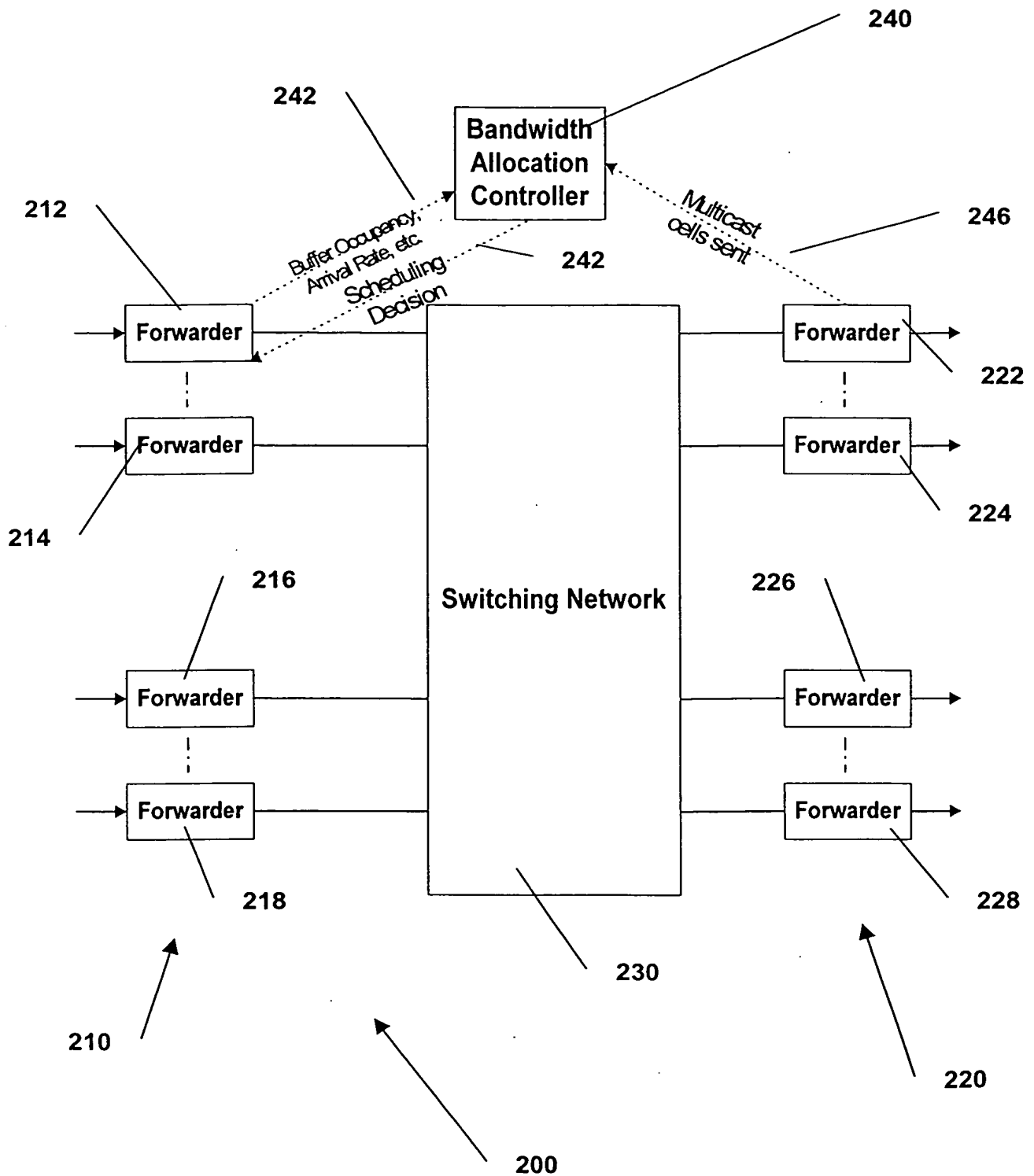
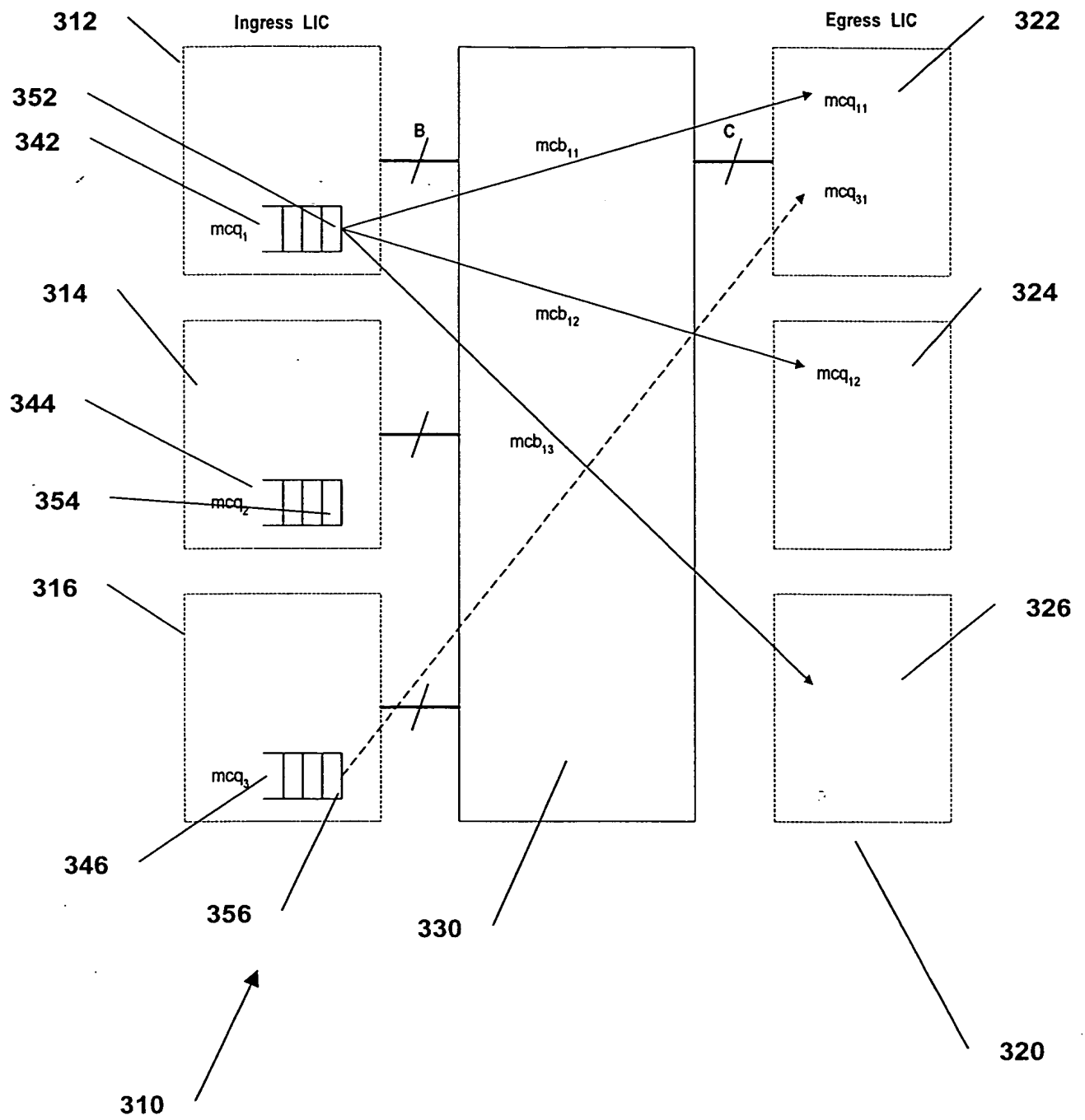


Fig. 1

THIS PAGE BLANK (USPTO)

**Fig. 2**

THIS PAGE BLANK (USPTO)

**Fig. 3**

THIS PAGE BLANK (USPTO)